(Thesis Synopsis - 5000 words)

Functions of bio-molecules deviating from canonical structures

Summary of work for the degree of Doctor of Philosophy (Ph.D.) in Chemistry (Theoretical)
from
Presidency University

Submitted by

Kanika Kole

Department of Physics of Complex Systems,
S. N. Bose National Centre for Basic Sciences,
Block- JD, Sector- III, Salt Lake, Kolkata - 700106, India
(Email: kanikakole0094@gmail.com / kanika1994@bose.res.in)

Ph.D. Registration Number: RC001-22RS205210349
Ph.D. Registration Date: 12th August, 2022

under the guidance of

Prof. Jaydeb Chakrabarti

Department of Physics of Complex Systems,
S. N. Bose National Centre for Basic Sciences,
Block- JD, Sector- III, Salt Lake, Kolkata - 700106, India

Jaydeb Chakrabarti 11/4/25

Signature of the supervisor
Prof. Jaydeb Chakrabarti
Senior Professor

Kanika Kole 11.04.25

Signature of the student
Kanika Kole
Senior Research Fellow

प्रो. जयदेब चक्रवर्ती / Prof. Jaydeb Chakrabarti
वरिष्ठ प्रोफेसर / Senior Professor
जटिल प्रणाली भौतिकी विभाग
Department of Physics of Complex Systems
सत्येंद्र नाथ बसु राष्ट्रीय मौलिक विज्ञान केंद्र
S. N. Bose National Centre for Basic Sciences
ब्लॉक - जे. डी., सेक्टर-III, सॉल्ट लेक, कोलकाता-106
Block-JD, Sector-III, Salt Lake, Kolkata-106

List of Publications

1. **Kanika Kole**; Aayatti Mallick Gupta and Jaydeb Chakrabarti; Conformational stability and order of Hoogsteen base pair induced by protein binding. *Biophys. Chem.* **2023**, *301*, 107079.
(https://doi.org/10.1016/j.bpc.2023.107079)

2. **Kanika Kole** and Jaydeb Chakrabarti; Binding of Homeodomain Proteins to DNA with Hoogsteen Base Pair. *J. Phys. Chem. B* **2025**, *129*, 1544-1554.
(https://doi.org/10.1021/acs.jpcb.4c08054)

3. **Kanika Kole**; Aayatti Mallick Gupta and Jaydeb Chakrabarti; The effect of protein binding on conformational stability and order of DNA duplex with Hoogsteen base pairing *AIP Conf. Proc.* **2025**, *3198*, 020005. (https://doi.org/10.1063/5.0248431)

Under preparation

1. **Kanika Kole** and Jaydeb Chakrabarti; Free energy landscape between Watson-Crick to Hoogsteen base pair transition.

2. **Kanika Kole** and Jaydeb Chakrabarti; Conformational stability of α-synuclein in presence of ZnO nanoparticle.

3. **Kanika Kole**; Abhik Ghosh Moulick and Jaydeb Chakrabarti; Coarse-grained model of proteins with structural informations.

Attending conferences

1. RARE 2025, held from March 9–12, 2025, at the Ramada by Wyndham, Khajuraho, Madhya Pradesh, India (poster presentation).

2. Stat. Mech. Meet Kolkata (SMMK) 2024, held on September 27–28, 2024, at the S. N. Bose National Centre for Basic Sciences, Kolkata, India (poster presentation).

3. The 46th Indian Biophysical Society Meeting, held from March 15–17, 2024, at TIFR Hyderabad and the University of Hyderabad, India (poster presentation).

4. Bose Fest 2024, held on February 22–24, 2024, at the S. N. Bose National Centre for Basic Sciences, Kolkata, India (talk presentation).

5. Nano Bio Interface (NB1) 2024, held on February 18-19, 2024, at Jawaharlal Nehru University, New Mehrauli Road, New Delhi 110067, India (poster presentation).

6. The 67th DAE Solid State Physics Symposium (DAE-SSPS), held from December 20–24, 2023, at the Gandhi Institute of Technology and Management, Visakhapatnam, Andhra Pradesh, India (poster presentation).

7. The Condensed Matter, Materials Science and Statistical Physics Symposium (CMMSSPS) 2023, held on April 28–29, 2023, at Presidency University, Kolkata, India (poster presentation).

1 Introduction

Non-canonical structures in biomolecules, such as DNA, RNA and proteins play essential roles in various cellular processes. In DNA, non-canonical Hoogsteen (HG) base pairing is often observed as an alternative to canonical Watson-Crick (WC) base pairing[1] (Fig. 1). Similarly, proteins like amyloid-beta (Aβ), α-synuclein (αS) lack a specified three-dimensional canonical structure under physiological conditions. Such proteins are known as intrinsically disordered proteins (IDPs)[2] (Fig. 2). Non-canonical biomolecules, although important in many functions, are yet not well understood in the literature.

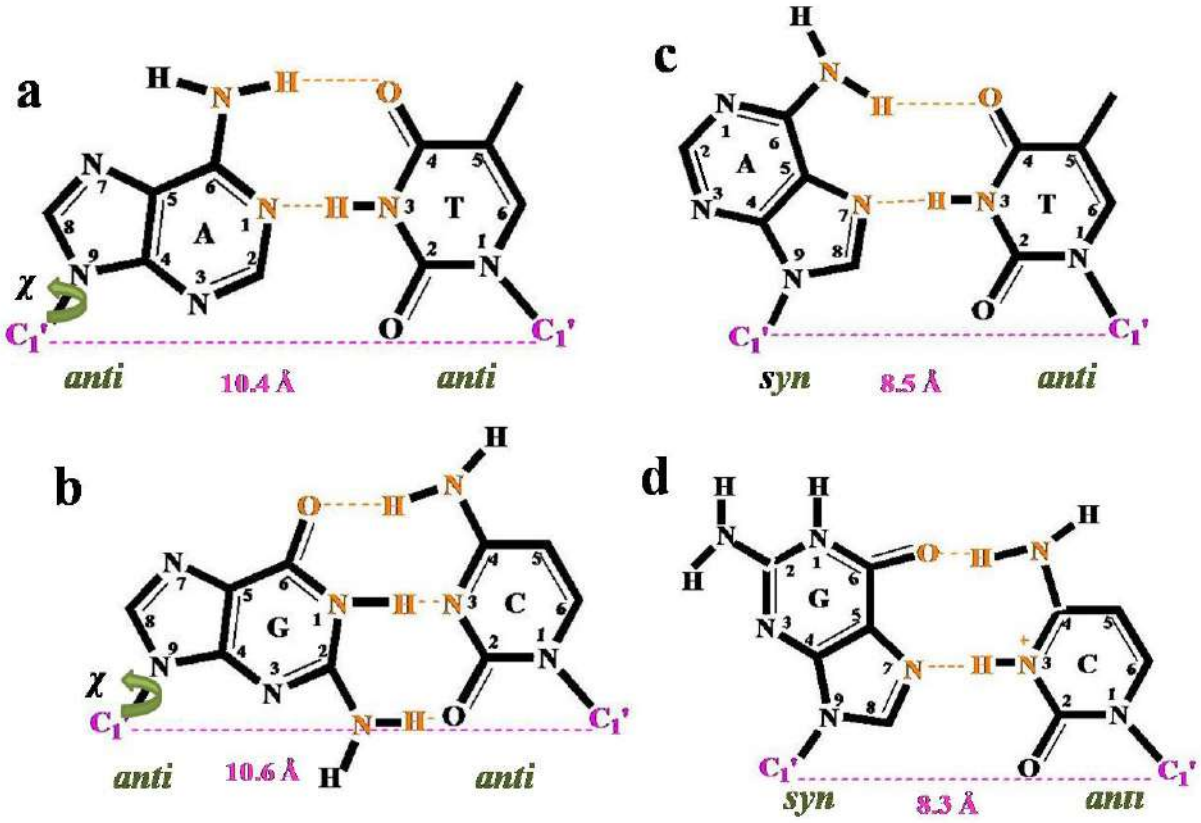

Figure 1: Chemical structure of (a) A-T and (b) G-C WC bp, (c) A-T and (d) G-C$^+$ HG bp. Heavy atoms involved in the WC and HG type H-bond are highlighted by orange. Average C1′-C1′ distance for each bp type is highlighted by pink. The χ angle is highlighted in green.

A specifying feature of DNA is the canonical WC base pairing for adenine(A)-thymine(T) and guanine(G)-cytosine(C). However, these bases can adopt an alternative base pairing mode known as HG base pairing, in which the purine base adopts a syn conformation instead of the anti conformation seen in WC pairs, leading to hydrogen bonding through the HG edge rather than the WC edge[3] (Fig. 1). This structural shift results in distinct base pair (bp) geometries

(Fig. 1), which have been observed in DNA duplexes under physiological conditions.[4] HG bps have significant biological implications, particularly in DNA-protein interactions,[5] gene regulation,[6] DNA repair[5] and drug-DNA interactions.[7,8] Experimental and earlier simulation studies have shown that HG bps are stabilized in DNA-protein complexes.[9] However, the microscopic mechanisms underlying the protein induced stability of HG bps have not yet been investigated.

A protein in its native form typically has a well-defined spatial distribution of the atoms, called globular protein. Instead of adopting a single well-defined conformation like globular proteins, IDPs exist as dynamic ensembles of conformations, allowing them to be highly flexible and functionally versatile.[10,11] This flexibility allows them to play crucial roles in cell signaling, regulation and molecular recognition.[11] IDPs are not inherently toxic. Sometimes, their dynamic and flexible nature make them prone to misfolding and aggregation leading to pathological conditions.[12,13]

Here we focus on understanding the microscopic description of proteins and DNA bp away from their canonical structure. We perform all atom molecular dynamics (MD) simulations to investigate the structural characteristics and conformational thermodynamics of WC and HG bp states in both naked and protein-DNA duplexes (PDB ID: 1K61), focusing on their interactions with homeodomain proteins.[14] We observe that the presence of proteins stabilized and ordered the HG bp as well as the whole DNA duplex in the 1K61 system with respect to the all WC bps containing naked DNA system. To examine how conformational stability and order at the protein-DNA interfaces alter the binding energy (B_E) of proteins[15] we compute the conformational stability and order using the conformational thermodynamics calculations[14–21] and evaluate the B_E of proteins using Umbrella Sampling (US) technique[22–24] for each DNA-protein system. We observe a strong correlation between conformational thermodynamics data at the interface and the B_E of protein with DNA.[15] This suggests that changes in conformational thermodynamics at the interface play a key role in protein-DNA binding. We also compute the relative free energy of the WC and HG base pairing modes in naked and protein-bound DNA systems using Well-Tempered Metadynamics (WT-MetaD)[25] technique. We use glycosidic angle, χ as the collective variable. Our results reveal that HG base pairing is ~ 3 kcal/mol less stable than WC base pairing in the absence of proteins but it becomes ~ 3.5 kcal/mol more stable in presence of proteins in this 1K61 system.

Now we consider proteins deviating from native structure. αS is a well known IDP and its

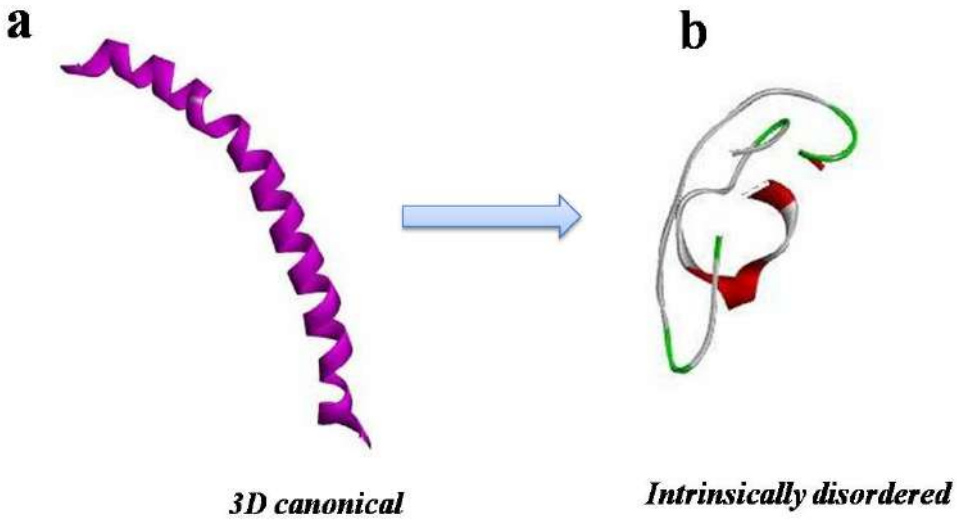

Figure 2: A cartoon representation of intrinsically disordered protein (IDP) in their (a) 3D canonical and (b) intrinsically disordered forms.

aggregation is a hallmark of Parkinson's disease.[26] Non-Amyloid-β Component (NAC) region (residues 61–95) of αS is the key driver of fibril formation in the Parkinson's disease. Recent studies suggest that ZnO nanoparticle (ZnONP) may alters the conformation of αS and reduce the rate of amyloid formation.[27] However, it has not been studied in detail how ZnONP affects on the conformational changes of αS. Computational approaches have provided extensive detail about some steps of the aggregation process of αS.[28] We use all atom MD simulations to investigate microscopically the effect of the ZnONP on the conformational changes of αS. The secondary structure analysis[29] shows that the helix percentage in the NACore (68GAVVTGVTAVA78) residues increases in the holo form compared to the apo form. Conformational thermodynamics[14–21] analysis reveal that the aggregation-prone NACore region of αS becomes stabilized in the presence of ZnONP. These results suggest that ZnONP may help to prevent αS aggregation by increasing the stability of the NACore residues.

Fully microscopic simulation calculations using the all atom simulation technique are quite expensive on IDP aggregation relevant to pathogenic behavior. This demands for coarse-grained (CG) description of systems. Several CG techniques have been reported in the literature. In most reported CG models for the protein in the literature, each amino acid is either represented as hydrophobic or polar entity which ignore protein backbone information.[30–34] Therefore, the main drawback of literature reported CG model is these kinds of model does not provide

any structural information. Motivated by this, we model the system using a CG approach,[35] representing the center of mass of each protein residue as a polymer bead and the oxygen atom of each water molecule as a solvent bead. Each bead is taken to have internal degrees of freedom corresponding to the dihedral angles. Then we simulate the system for a longer time using the Monte Carlo (MC)[36] method. This CG method incorporates all the non-bonded, bonded interactions, along with dihedral angle coupling information derived from equilibrated all atom data. We apply our CG model to the small proteins GB3 (56 residues, PDB ID: 2OED). Our CG simulation data show strong structural agreement with both the crystal structure and all atom MD data. We now apply this CG approach to the αS protein to examine how ZnONP reduces the fluctuation in the NACore residues of αS. Our CG approach reproduces the all atom data almost well.

The synopsis contains the following sections: Sec. 2 describes the conformational stability and order of the HG bp in presence of homeodomain proteins. Sec. 3 contains the correlation between the binding energy of homeodomain protein with DNA and conformational thermodynamics data at the protein-DNA interface considering both the HG-DNA and WC-DNA protein systems. The free energy landscape between Watson-Crick and Hoogsteen base pairing transitions are described in sec. 4. The effect of ZnO nanoparticle on the conformational changes of the α-synuclein using all atom molecular dynamics simulation is discussed in the sec. 5. Coarse-grained Monte Carlo simulation of the protein is described in the sec. 6. We conclude the synopsis in sec. 7.

2 Conformational Stability and Order of Hoogsteen Base Pair Induced by Protein Binding

HG bps are commonly observed in DNA-protein complexes and are stabilized through interactions with proteins.[9] HG bp occurs in regions of DNA that are highly distorted by bound protein.[37] The X-ray structure of TATA-binding protein (TBP) bound to DNA reveals a protonated G-C^+ HG bp in the region of DNA under winding and intercalation by TBP's phenylalanine side chain.[38] Furthermore, the X-ray crystal structure of DNA in contact with the p53 tumour suppressor proteins exhibits two consecutive A-T HG bps. Wolberger and co-workers report a single A-T (A_7 -T_{26}) HG bp within an undistorted B-form DNA duplex (PDB ID: 1K61) in

the X-ray structure of MATα2 homeodomain bound to DNA.[9] The crystal structure contains 21 fragments of duplex DNA attached to four MATα2 homeodomains. The DNA shows two binding sites for the four MATα2 homeodomains, two homeodomains, α2A and α2B bind at the specific binding. sites, while the other two homeodomains, α2C and α2D bind to DNA non-specifically. Earlier MD simulation studies show that the formation of HG bp in the 1K61 system is more strongly influenced by the non-specifically bound protein α2D.[9] However, there has yet not been any attempt to understand the microscopic mechanism underlying the protein induced stability of the HG bp.

We use conformational thermodynamics calculation to investigate the stability of the HG bp in this system at the molecular level using conventional MD trajectories.[14] The changes in conformational stability and order in bio-molecular systems are given in terms of conformational free energy and entropy changes derived from the fluctuations of microscopic conformational variables. Negative changes in conformational free energy and entropy indicate structural stability and ordering respectively in a given structure with respect to a reference structure, while positive changes indicate conformational destabilization and disorder respectively. We use the microscopic conformational variables of DNA are (i) inter-bp step parameters: shift (D_x), slide (D_y), rise (D_z), tilt (τ), roll (ρ) and twist (ω) (ii) intra-bp parameters: stagger (S_x), shear (S_y), stretch (S_z), buckle (κ), propeller (π) and opening (σ) (iii) sugar-phosphate and sugar-base backbone torsion angles: α, β, γ, δ, ϵ, ζ and χ, and (iv) sugar-puckers: $\nu0$, $\nu1$, $\nu2$, $\nu3$ and $\nu4$.

We perform conventional MD simulations on three different systems: (1) WC, a model system of DNA, (2) HG, HG bp containing DNA (PDB ID: 1K61), and (3) HGP, protein-bound HG bp containing DNA (PDB ID: 1K61). The equilibrated snapshots at 500 ns time span of the WC, HG, and HGP systems are shown in Figs. 3 (a), (b) and (c), respectively. First, we check the HG base pairing criteria and compare them with the literature values. The A_7 -T_{26} bp in the HG and HGP systems maintains HG base pairing criteria throughout the equilibrated trajectory. In the HGP system, we identify the residues of the protein that form strong contact with A_7 and T_{26} bases. We note that protein α2D residues ARG132, GLY133 and GLN175 create interface with base A_7 of DNA. These three residues form hydrogen-bonds (H-bonds) with A_7 at the interface. Moreover, residue ARG132 is involved in forming salt bridge as well as electrostatic contact between the N atom of ARG132 and the O1P atom of A_7. On the other hand, the specific protein α2B residues HIS134, PHE136, and TRP179 create H-bonds with T_{26}.

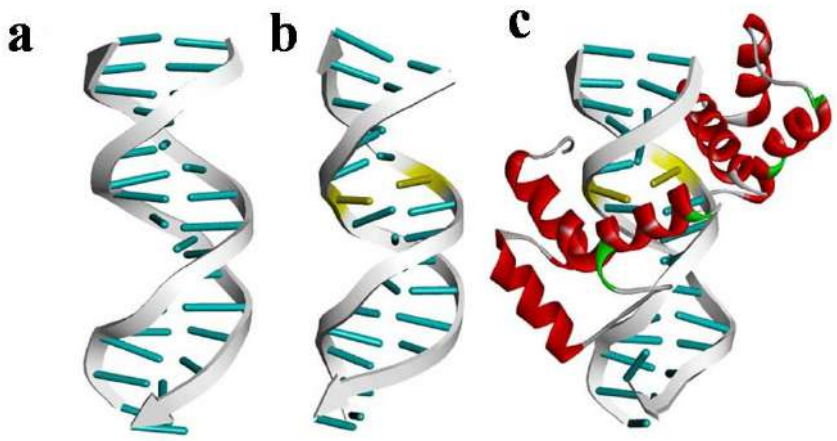

Figure 3: Snapshot at 500 ns of (b) WC (c) HG (d) HGP systems. HG bp is shown in yellow.

Therefore, the protein α2D interacts more strongly with A_7 compared to α2B with T_{26}.

The conformational fluctuations are quantified in terms of the distribution, $H^S(\theta_R)$, (S=WC, HG and HGP) of a microscopic conformational variable, θ pertaining to base, R over the equilibrated trajectory. Each peak of the histogram represents the most probable value for the given variable. The larger width of the histogram peaks denotes more conformational flexibility. The HG system shows multi-peaked or broad distribution in most of the microscopic conformational degrees of freedom compared to the WC system for A_7 - T_{26} bp. This suggests that HG bp is more flexible than WC bp in naked DNA. On the other hand, in the majority of cases, the sharper peaks of the HGP complex indicate that the presence of proteins reduce the flexibility of the HG bp. To compute quantitatively, we perform conformational thermodynamics calculations.

We compute conformational thermodynamics data for each microscopic degrees of freedom : (i) in the HG system and (ii) in the protein-bound HGP system, both with respect to the WC system. We calculate the total changes in conformational free energy and entropy for entire DNA systems. The changes in the distribution of the microscopic conformation variables lead to the changes in conformational free energy and entropy of DNA in HG and HGP systems with reference to the WC system. By summing the changes in individual microscopic degrees of freedom, we determine the overall conformational thermodynamic changes. We compute the changes in conformational free energy and entropy, $\Delta G^{HG,inter}$ and $T\Delta S^{HG,inter}$ for inter-bp parameters, $\Delta G^{HG,intra}$ and $T\Delta S^{HG,intra}$ for intra-bp parameters, $\Delta G_d^{HG,SP}$, $\Delta G_c^{HG,SP}$, $T\Delta S_d^{HG,SP}$ and $T\Delta S_c^{HG,SP}$ for sugar-base and sugar-phosphate backbone torsion angles, $\Delta G_d^{HG,S}$, $\Delta G_c^{HG,S}$,

Table 1: The changes (kJ/mol) in conformational thermodynamics of HG and HGP systems in comparison to the WC system.

$\Delta G^{HG,inter}$	$T\Delta S^{HG,inter}$	$\Delta G^{HG,intra}$	$T\Delta S^{HG,intra}$	$\Delta G_d^{HG,SP}$	$\Delta G_c^{HG,SP}$	$T\Delta S_d^{HG,SP}$	$T\Delta S_c^{HG,SP}$	$\Delta G_d^{HG,S}$	$\Delta G_c^{HG,S}$	$T\Delta S_d^{HG,S}$	$T\Delta S_c^{HG,S}$	ΔG_{total}^{HG}	$T\Delta S_{total}^{HG}$
1.57	0.73	0.20	1.62	0.17	-0.61	-5.27	-2.57	0.02	-0.03	-1.1	2.78	1.32	-3.81
$\Delta G^{HGP,inter}$	$T\Delta S^{HGP,inter}$	$\Delta G^{HGP,intra}$	$T\Delta S^{HGP,intra}$	$\Delta G_d^{HGP,SP}$	$\Delta G_c^{HGP,SP}$	$T\Delta S_d^{HGP,SP}$	$T\Delta S_c^{HGP,SP}$	$\Delta G_d^{HGP,S}$	$\Delta G_c^{HGP,S}$	$T\Delta S_d^{HGP,S}$	$T\Delta S_c^{HGP,S}$	ΔG_{total}^{HGP}	$T\Delta S_{total}^{HGP}$
-2.41	-16.16	-0.60	-2.25	-5.69	-2.56	-21.74	-12.31	-8.74	-5.24	-13.74	-7.42	-25.23	-73.62

$T\Delta S_d^{HG,S}$ and $T\Delta S_c^{HG,S}$ due to sugar-puckers, in the HG system with respect to the WC system summing over all microscopic degrees of freedom. Here, ″d″ refers to the 5′-3′ strand and ″c″ refers 3′-5′ strand. The data are shown in Table 1. We find that most of the changes in conformational free energy and entropy are insignificant compared to the thermal energy ($\approx$ 2.5 kJ/mol) at room temperature. Only the conformational entropy, $T\Delta S_d^{HG,SP}$ and $T\Delta S_c^{HG,SP}$ show ordering due to sugar-base and sugar-phosphate torsion angles. Similarly, we show the analogous conformational thermodynamic quantities $\Delta G^{HGP,inter}$ and $T\Delta S^{HGP,inter}$, $\Delta G^{HGP,intra}$ and $T\Delta S^{HGP,intra}$, $\Delta G_d^{HGP,SP}$, $\Delta G_c^{HGP,SP}$, $T\Delta S_d^{HGP,SP}$ and $T\Delta S_c^{HGP,SP}$, $\Delta G_d^{HGP,S}$, $\Delta G_c^{HGP,S}$, $T\Delta S_d^{HGP,S}$ and $T\Delta S_c^{HGP,S}$ in the HGP system with respect to the WC system in Table 1. We observe significant stabilization in $\Delta G_d^{HGP,SP}$, $\Delta G_d^{HGP,S}$ and $\Delta G_c^{HGP,S}$ and ordering in $T\Delta S^{HGP,inter}$, $T\Delta S_d^{HGP,S}$, $T\Delta S_c^{HGP,S}$, $T\Delta S_d^{HGP,S}$ and $\Delta G_c^{HGP,S}$ in the HGP system with respect to the WC system.

Finally, we compute ΔG_{total}^{HG} and $T\Delta S_{total}^{HG}$, summing all the individual components for the entire HG system with respect to the WC system are 1.33 kJ/mol and -3.81 kJ/mol, respectively (Table 1). Similarly, we compute ΔG_{total}^{HGP} and $T\Delta S_{total}^{HGP}$ for the whole HGP system in comparison to the WC system are -25.23 kJ/mol and -73.62 kJ/mol, respectively (Table 1.). Overall, the total changes of conformational thermodynamics of the entire DNA system, we observe that the HG system is slightly destabilized but ordered than the WC system, while the HGP system is significantly stabilized and ordered in comparison to the WC system. The inter-bp step parameters are the primary factors to destabilize the HG system compared to the WC system, while sugar-puckers are the main factors to stabilize the HGP system over the WC system. We observe in Table 1. that for the 5′-3′ strand, $\Delta G_d^{HGP,SP}$, $\Delta G_d^{HGP,S}$, $T\Delta S_d^{HGP,SP}$ and $T\Delta S_d^{HGP,S}$ are more negative compared to $\Delta G_c^{HGP,SP}$, $\Delta G_c^{HGP,S}$, $T\Delta S_c^{HGP,SP}$ and $T\Delta S_c^{HGP,S}$, respectively, of the complementary 3′-5′ strand. This shows protein α2D has a greater impact on stabilizing and ordering the HGP system than protein α2B. We observe that the HG bp region in the HGP system stabilized and ordered more than the WC system. Fig. 4 color-coded cartoon presentation of the HG and HGP systems with respect to the WC system. The microscopic

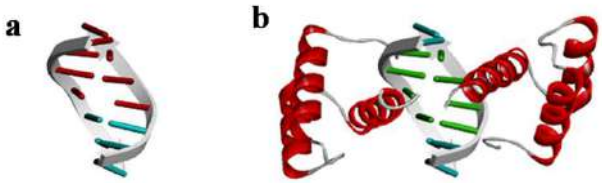

Figure 4: Color-coded cartoon presentation of the (a) HG and (b) HGP systems, with respect to the WC system. Compared to the WC system, slightly disordered or ordered bps of the HG system are shown in red, whereas significantly stabilized and ordered bps of the HGP system are shown in green.

degrees of freedom D_x, χ, δ, ϵ, ζ, $\nu 0$, $\nu 1$, $\nu 2$, $\nu 3$ and $\nu 4$ play a crucial role in this stabilization. This gives a more complete microscopic picture of the stability at the HG bp induced by the bound protein.

3 Binding of Homeodomain Proteins to DNA with Hoogsteen Base Pair

The B_E of proteins with DNA is crucial as it determines the strength, specificity, and stability of the interaction, directly influencing gene regulation, DNA repair, and other biological functions.[39,40] The B_E of proteins to WC-DNA has been reported in both experimental studies[41] and simulations.[42,43] However, the B_E of proteins binding to HG-DNA has not been reported so far. Using US simulation method,[22–24] we compute the B_E of both proteins in the WCP system, where all bps follow the WC type and compare the results with the wild-type HGP system. Mutations can significantly alter protein-DNA interactions by affecting B_E. To examine these effects, we introduce mutations in the α2D protein of the HGP system. In one case, we replace the basic ARG132 residue with a neutral ASN132 (ARG132ASN), forming the NMUTP system. In another case, we substitute ARG132 with an acidic ASP132 (ARG132ASP), creating the AMUTP system. Our primary objective is to determine how interfacial stability and order influence the B_E in these protein-DNA systems.[15]

We compute the B_E of the protein with the rest of the system using ζ as the reaction coordinate that represents the center of mass distance between the protein and the rest of the system. First, we consider the B_E calculation for the α2D protein with remaining portion of the HGP system. We run three independent US simulations to compute the average PMF curve. The

average PMF curve and associated errors are computed using the results of the three independent US datasets as shown in Fig. 5 (a). The inset figure in Fig. 5 (a) shows the magnified values of $V_{PMF}(\zeta)$ for low ζ. To calculate the B_E, we take the difference between the highest and lowest values of the average PMF curve.[44–50] We compute the B_E to be approximately -36.57 $\pm$ 0.78 kcal/mol for the α2D protein with rest part of the HGP system. The similar protocol is followed to compute the B_E for the other cases. Similarly, we compute the B_E of the α2B protein with the rest part of the HGP system. The B_E of α2B is around -37.00 $\pm$ 0.56 kcal/mol as shown in Fig. 5 (b). Therefore, in the HGP system, the B_E of the non-specific protein, α2D, and the specific protein, α2B, with HG-DNA in the presence of the other protein are quite close.

Then we check the B_E of both the α2D and α2B proteins with rest part of the WCP system. Figs. 5 (c) and (d) show that the B_E of the α2D and α2B proteins in the WCP system decreases compared to the HGP system. Therefore, both the proteins show stronger binding in the HGP system than the WCP system.

Next we consider a situation where HG-DNA is bound to one of the proteins, either α2D or α2B in the absence of the other protein. Then we compute the B_E of the α2D and α2B protein. From Figs. 5 (e) and (f) in the absence of other protein, the B_E of the α2D and α2B proteins with HG-DNA decrease, suggesting α2D and α2B bind cooperatively to HG-DNA in the HGP system.

Then we compute the B_E of the α2D and α2B proteins with the rest part of the NMUTP and AMUTP systems (Figs. 5 (g-j)). In both the NMUTP and AMUTP systems, the B_E of α2D decreases, while the B_E of α2B decreases in the NMUTP system but increases in the AMUTP system as shown in Figs. 5 (g-j) . We then compare the B_E across all cases with interfacial interactions between the protein and DNA, observing that stronger interfacial interactions correspond to higher B_E.

We further quantify the relation between the B_E and interfacial stability and order. We compute conformational free energy and entropy from the fluctuations of the microscopic conformation variables[14–21] at the interface between protein and DNA . For the protein, we consider the conformational variables ϕ, ψ and $\chi 1$ dihedral angles. The changes in the distribution of the microscopic degrees of freedom lead to the changes in the conformational free energy and entropy. Next, we compute the conformational thermodynamics data by adding the data of the DNA bases as in Ref. 15 and protein residues as in Ref. 22 at the interfaces. Fig. 6(a) shows

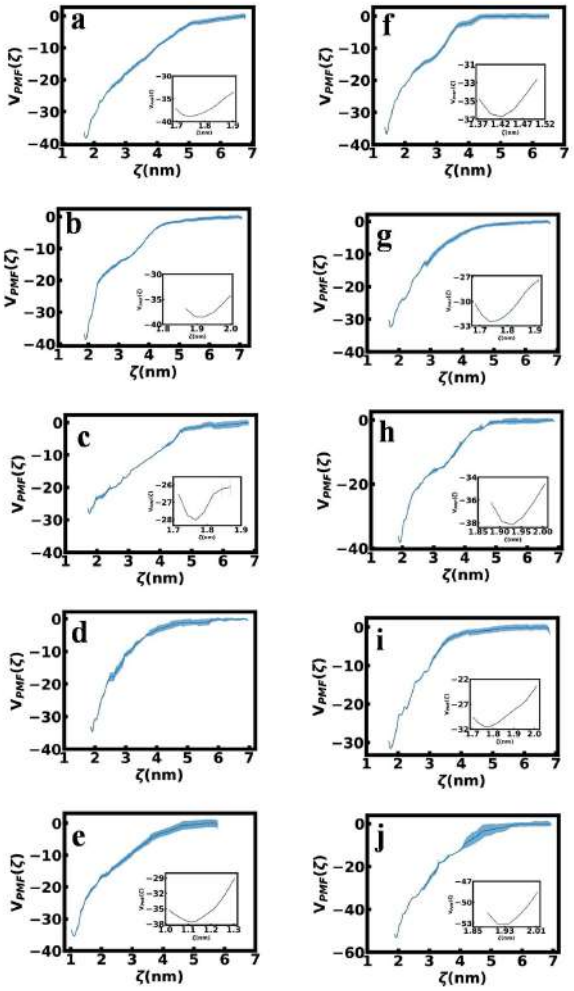

Figure 5: Average PMF curves with error bars for (a) α2D and (b) α2B proteins, respectively with the HG-DNA duplex in the presence of other protein obtained from the US method for the HGP system. Similarly, average PMF curves with error bars for (c) α2D and (d) α2B proteins, respectively, with the remaining part of the WCP system. (e) Average PMF curve with error bars for α2D protein with HG-DNA duplex in the absence of α2B protein and (f) average PMF curve with error bars for α2B protein with HG-DNA duplex in the absence of α2D protein, respectively. Average PMF curves with error bars for (g) α2D and (h) α2B proteins, respectively with the rest of the NMUTP system. (i) Average PMF curves with error bars for α2D and (j) α2B proteins, respectively, with the remaining part of the AMUTP system. The inset figures in each of the figure show magnified view of $V_{PMF}(\zeta)$ at low ζ values.

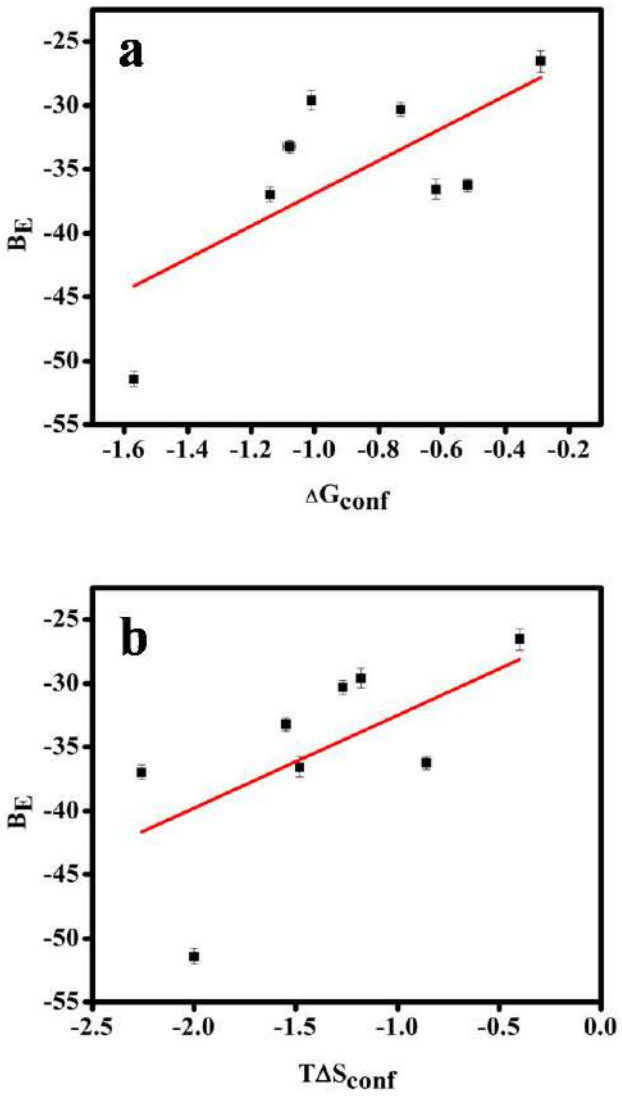

Figure 6: (a) Correlation plot between conformational free energy, ΔG_{conf}, of interface forming residues and binding energy B_E with error bars of proteins, $\alpha 2D$ and $\alpha 2B$. (b) Correlation plot between conformational entropy, $T\Delta S_{conf}$, with error of interface forming residues and binding energy B_E with error bars of $\alpha 2D$ and $\alpha 2B$ proteins. All data in kcal/mol. Black square symbols represent the data points and the red dotted line represents linear fitting of the data points.

the linear correlation between conformational free energy, ΔG_{conf}, at the interface and binding energy, B_E, of proteins for different DNA-protein systems in Fig. 6(a). The Pearson correlation coefficient[51–53] r_{xy} is 0.72, suggesting strong correlation between stability of the interface and B_E. The positive correlation between conformational entropy, $T\Delta S_{conf}$ at the interface and B_E of proteins is shown in Fig. 6(b). Here, the Pearson correlation coefficient[51–53] is 0.66, indicating moderate positive correlation between $T\Delta S_{conf}$ and B_E. Therefore, Figs. 6(a) and (b)

show the B_E of proteins increases with an increase in the conformational stability and order of protein residues and DNA bases at the interface. The strong correlation between conformational thermodynamics data at the interface and the binding energy between the homeodomain protein and DNA in Fig. 6, suggests that changes in conformational thermodynamics at the interface play a crucial role in protein-DNA binding. The correlation between conformational thermodynamics and B_E can be experimentally verified. Such a plot could be valuable for calibrating of protein binding to DNA.

4 Free Energy Landscape between Watson-Crick to Hoogsteen Base Pairing Transitions

The inherent flexibility of double-helical DNA enables spontaneous conformational fluctuations, known as DNA breathing, which allow bps to adopt geometries beyond conventional WC pairing.[54] One example of DNA breathing is the thermal transition from WC to HG base pairing.[54] Such bp transitions can modify DNA's local and global geometry, facilitating enzyme recognition and DNA processing.[54]

Experimental NMR relaxation studies have reported that the transition between WC and HG base pairing occurs on a timescale of 50–250 ms, which exceeds the capabilities of conventional MD simulations with currently available computing facilities.[55] Therefore, the transition between these two conformations can only be studied using enhanced sampling methods, such as umbrella sampling,[56] metadynamics,[57,58] and meta-eABF.[59] Previous literature reported simulation studies[1,54] based on two CVs, the glycosidic angle (χ) and the pseudodihedral angle (ϕ), show that bp transitions at small ϕ angle values ($\sim 0°$) are energetically more favorable. However, the free energy landscape for WC to HG bp transition in presence of proteins has not been reported.

We primarily focus on the lowest energy transition pathway. We use the WT-MetaD technique[60] to compute free energy profiles for WC and HG bp breathing fluctuations in both naked DNA and protein-DNA complexes, with the glycosidic angle (χ) as the CV. We take the amber14SB_OL15 force field for our calculations. We consider four systems in our calculations: WC (naked DNA with all WC bps), HG (naked DNA with one HG bp and the rest are WC bps), WCP (protein-DNA complex with all WC bps) and HGP (protein-DNA complex with

one HG bp and the rest are WC bps). Our focus is on the DNA breathing fluctuations at the A_7-T_{26} bp in each system. In the WC system, we observe that the energy barrier for the WC to HG bp transition is approximately 11 kcal/mol, with the WC bp being about 3.2 kcal/mol more stable than the HG bp (Fig. 7 (a)). These results are consistent with experimental data[61] and previous simulation studies.[1,54] In the WCP system, where proteins are present, the energy barrier for the WC to HG transition is reduced to around 8 kcal/mol, with the HG bp being approximately 3.3 kcal/mol more stable than the WC bp Fig. 7 (b). For the HG to WC bp transition in the HG system, the energy barrier is around 7 kcal/mol, with the WC bp being about 3.2 kcal/mol more stable than the HG bp Fig. 7 (c). However, in the presence of proteins in the HGP system, the energy barrier increases to approximately 12 kcal/mol for the HG to WC bp transition, with the HG bp being about 3.5 kcal/mol more stable than the WC bp Fig. 7 (d). We compare our WC system results with those reported in the literature using other amber force fields (amber14sb_parmbsc1, amber99sb_parmbsc0) and find the results to be similar. We expect that this study will provide a robust computational framework to complement and guide ongoing DNA breathing experiments in both naked and protein-bound DNA systems.

5 Conformational stability of α-synuclein in presence of ZnO nanoparticle

The self-aggregation of proteins into amyloid fibrils is associated with several neurodegenerative diseases, including Parkinson's disease (PD), Alzheimer's disease and Huntington's disease.[62–64] The fibrillation of αS is mainly related to PD.[65–67] αS is abundant in neuronal cells and plays a crucial role in synaptic transmission by facilitating synaptic vesicle release and contributing to synaptic membrane recycling.[68] The amino acid sequence of αS is typically divided into three regions based on its physicochemical properties: the amphipathic N-terminal domain (residues 1–60), the highly hydrophobic central domain (residues 61–95), known as NAC, and the acidic C-terminal domain (residues 96–140).[69] The NAC domain of αS is the key region responsible for protein self-aggregation and the conformational changes extending from the random coil to the β-sheet secondary structure.[70]

A recent experimental study shows that the presence of the ZnONP interface favors the formation of a stable αS-nanoparticle complex, leading to a relatively lower fraction of free αS

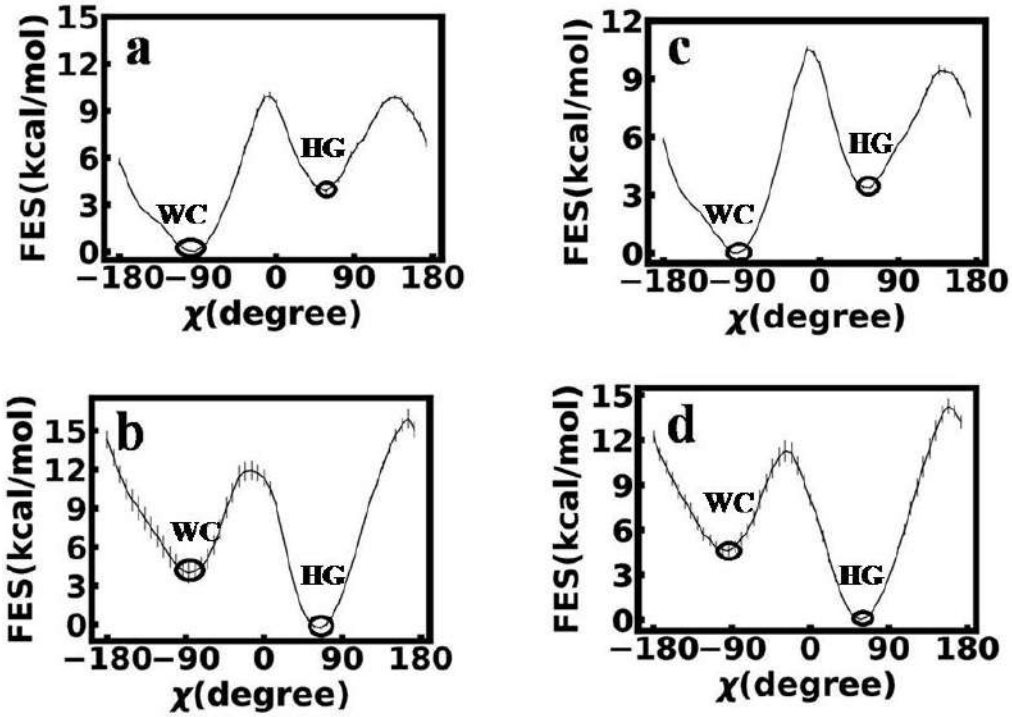

Figure 7: Free energy profile on χ for (a) the WC to HG bp transition in the WC system, (b) the WC to HG bp transition in the WCP system, (c) the HG to WC bp transition in the HG system and (d) the HG to WC bp transition in the HGP system.

monomers undergoing fibrillation.[27] Therefore, αS in complex with ZnONP exhibits significantly lower cytotoxicity against IMR32 and THP-1 cells in vitro compared to fresh αS.[27] However, how ZnONP inhibits fibril formation remains unclear. Further, a recent simulation study[69] shows that within a non-amyloid β component (NAC) spanning from the 61st to the 95th residue of α-synuclein, an 11-residue segment called NACore (68GAVVTGVTAVA78) is an essential region for both fibril formation and cytotoxicity.

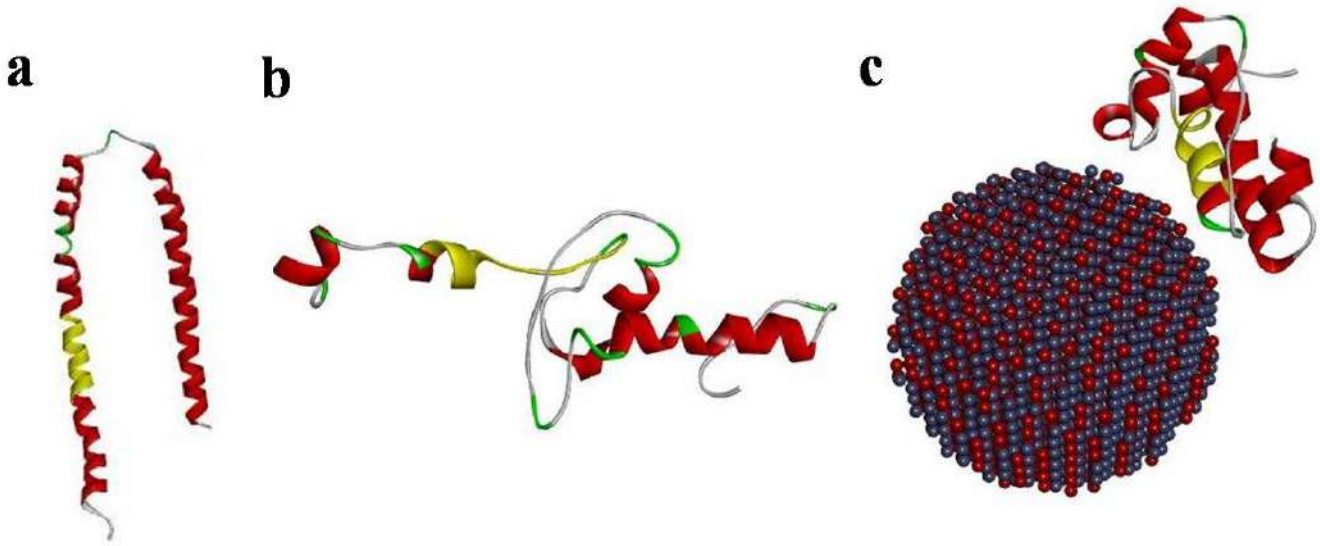

Figure 8: (a) Crystal structure of αS. Equilibrated snapshots of (b) apo and (c) holo systems, respectively. The NACore sequence (68GAVVTGVTAVA78) is highlighted in yellow.

We primarily focus in our studies on how the conformational stability and order of NACore residues in αS change in the presence of ZnONP. We perform all atom MD simulation on (1) α-synuclein (PDB ID: 1XQ8) (apo) and (2) α-synuclein-ZnONP (holo) systems, for 1 μs. We utilize a 4 nm diameter ZnONP. Fig. 8 (a) represents the crystal structure (PDB id-1XQ8). Figs. 8 (b) and (c) show equilibrated snapshots of apo and holo systems, respectively. We identify the interface forming residues of αS with ZnONP. The N-terminal domain of αS, rich in lysine residues, which is a part of the consensus hexameric (KTKEGV) motif, forms strong interfacial interaction with the ZnONP, consistent with experimental observations.[27] αS interacts with ZnONP through electrostatic interactions, salt bridges, and H-bonds. Secondary structure analysis reveals that the helix percentage increases in the NACore of the holo form with respect to the apo form as shown in Figs. 9 (a) and (b). Conformational thermodynamics calculations (Figs. 9 (c) and (d)) show that all the residues except one of NACore are stabilized

and ordered in the holo form with respect to the apo form. Using machine learning techniques,[71] we observe that the most essential coordinate responsible for αS's conformational changes shifts from the NAC region to the N-terminal region in the holo form. Thus, our all atom MD simulations data reveal that the ZnONP enhances the stability of NACore residues of αS in their helix from.

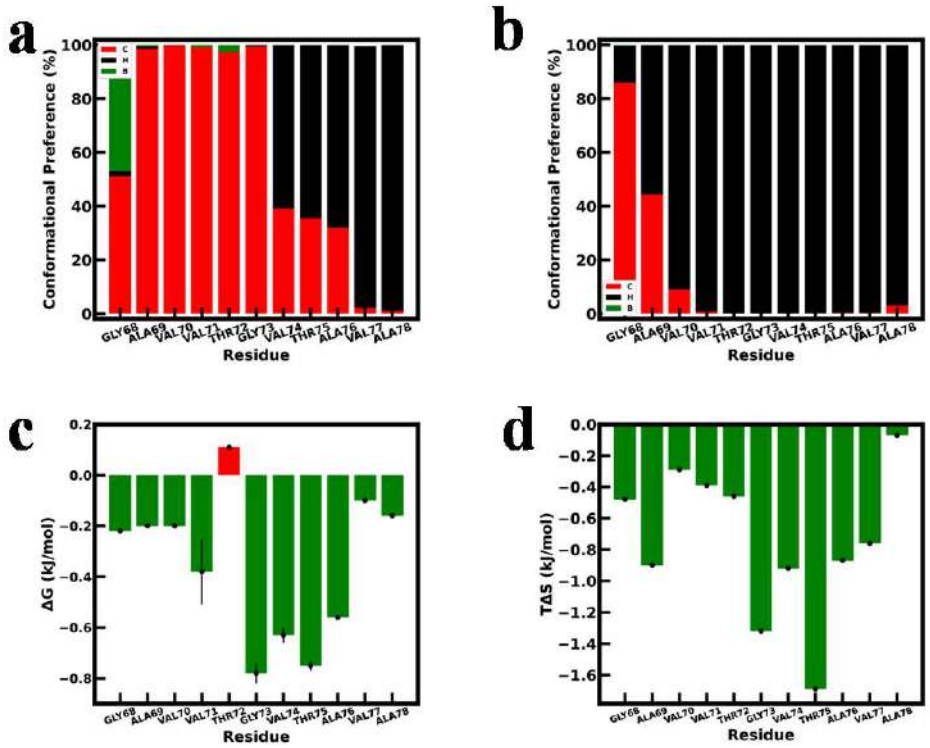

Figure 9: Conformational preference of NACore residues in (a) apo and (b) holo forms. Helices are shown in black, unstructured regions in red and β-sheets in green. (c) Conformational free energy and (d) conformational entropy of NACore αS residues in the holo form with respect to the apo form. Conformationally stabilized or ordered residues are in green and destabilized or disordered residues are in red.

6 Coarse-grained model for protein structure

A full atomic description of protein aggregation is computationally challenging. Reducing the representation from an all atom model to a CG model is essential for studying biomolecular systems with a large number of molecules over long time scales.[72,73] The main purpose of CG studies is to reduce the number of degrees of freedom, making it useful for studying large-scale conformational changes, protein-protein interactions and folding processes.[74] However,

the information on the dihedral angles are lost in the currently available models. To accurately capture the aggregation process, it is crucial to preserve secondary structure and dihedral angle information. Therefore, we build a simple polymer model to make CG description of protein with dihedral angle information.

In our CG model, we model the center of mass of residue of the protein as a polymer bead and the oxygen atom of water as the solvent bead. We set the bead diameter from the minimum of the negative logarithm of the radius of gyration distribution of the protein residues in all atom simulations. Polymer beads interact through both bonded and non-bonded interactions. All bonded and non-bonded interaction parameters are derived from all atom MD simulation data. Bonded interactions are primarily governed by the stretching of two consecutive beads and the bending of three consecutive beads, both follow the harmonic potential form. The bonded force constant is computed from the second order derivative at the minimum of the negative logarithm of the bond distance distribution between the center of mass of amino acid residues in all atom simulation data. Similarly, the angle force constant is also computed using the second order derivative at the minimum of the negative logarithm of the distribution of the bond-angle between center of mass of the residues. Non-bonded interactions are two types: the Lennard-Jones (L-J) 12 repulsive potential and screened Coulomb interactions. The L-J 12 interaction parameter, along with the constant and Debye length of the screened Coulomb potential are calculated from all atom MD simulation data of the protein. The bead-bead and solvent-bead interaction potential forms are computed to fit the density function of the solvent around the bead particles derived from all atom MD data. Each polymer bead is assigned with two additional degrees of freedom to represent two dihedral angles. Coupling between the dihedral in the model is given using free energy landscape obtained from all atom simulation data. For solvent-solvent bead interactions, only the non-bonded L-J 12-6 potential is considered. The interaction parameter is the minimum value of the negative logarithm of the solvent-solvent radial distribution function derived from all atom MD simulation data.

We use the MC method to generate conformations based on model interactions while maintaining the canonical (NVT) ensemble. When constructing MC moves, we calculate the energy cost arising from various interactions, including bonded, non-bonded and solvent contributions. We also change the dihedral degrees of freedom and compute the associated energy cost using a potential derived from all atom simulations through linear interpolation.

We apply our model to GB3 protein (PDB ID: 2OED).[75] GB3 is a small globular protein consisting of 56 amino acid residues. It serves as a model system for computational studies due to its well defined structure and compact size. We derive all the CG force field parameters from equilibrated trajectory in all atom MD of GB3 and perform the CG MC simulations. We perform 1,00,000 number of MC steps out of which first 30,000 MC steps are discarded for equilibration, based on potential energy of the system. The snapshots of the GB3 protein from the MC simulation are shown in in Figs. 10 (a) and (b). Different quantities like density profile of solvent around solvophilic and solvophobic bead, dihedral angle are averaged for configurations of multiple independent equilibrated trajectories. Figs. 11 (a) and (b) show the solvent distribution around solvophilic and solvophobic beads in the CG simulation and all atom simulation, respectively. The distributions show CG data matches well with the all atom MD simulation data. Next, we compare the structural elements of each residue from the CG simulation with those from the initial crystal structure and all atom simulation data. We find that the secondary structure elements of $\sim 75\%$ of residues in the CG simulation remain same with all atom MD simulation data, while $\sim 72\%$ match the crystal structure data.

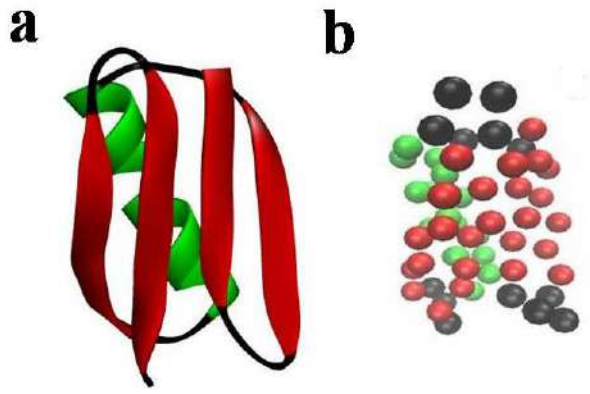

Figure 10: The snapshots of the GB3 protein in (a) crystal structure and (b) coarse grained (CG) simulations. Helices are shown in green, β-sheets in red and unstructured (other than helices or sheets i.e. loop/coil/turn/bend) regions in black.

We apply similar coarse graining technique for αS-ZnONP complex. We calculate the equilibrium fluctuation of αS both with and without ZnONP. In case of system containing αS and ZnONP, we keep fixed the binding residues as observed in all atom simulations. We find that ψ degrees of freedom in the NACore region gets stabilized compared with those without ZnONP. This is agreement with all atom data.

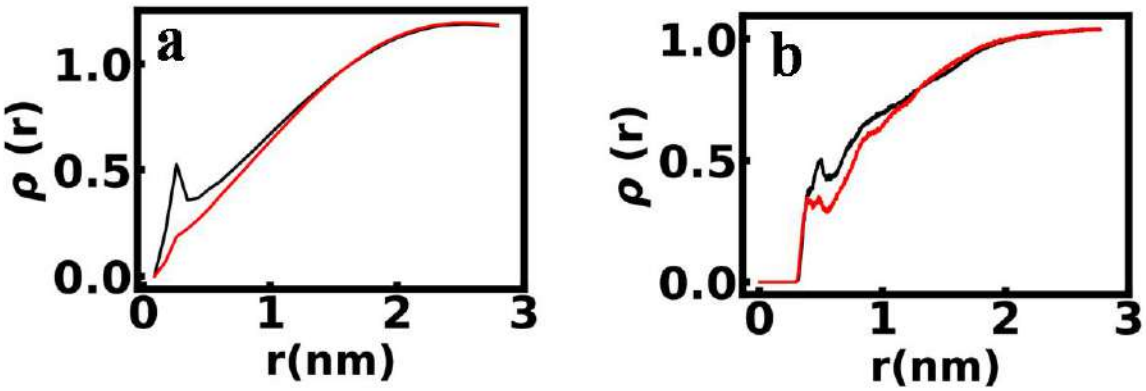

Figure 11: Solvent distribution around solvophilic and solvophobic beads in (a) CG simulation and (b) all atom MD simulation, with solvophilic beads shown in black and solvophobic beads in red.

7 Conclusion

In conclusion, we investigate the conformational stability and order of the HG bp containing duplex DNA in the presence of the homeodomain proteins with respect to the fully WC bps containing naked DNA. Additionally, we explore how the conformational stability and order at the protein-DNA interface influences the B_E of the homeodomains in the HG-DNA-protein systems. We compute the free energy differences between WC and HG bps, as well as the energy barrier for bp transformation, in both naked DNA and DNA-protein systems. Our findings provide valuable insights for designing stable HG-DNA–homeodomain protein systems, which may have significant implications in DNA replication, recognition, damage repair and other cellular processes. We also check the conformational stability of the IDP, αS in the presence of the ZnONP. Understanding the stability and interactions of the αS in presence of the ZnONP provides important revelations into their biological significance and potential applications in biomedical research. We develop a coarse-grained protein model to capture accurate structural information over longer time scales, which may be useful for studying the aggregation phenomena of IDPs.

References

[1] Yang, C.; Kim, E.; Pak, Y. Free energy landscape and transition pathways from Watson–Crick to Hoogsteen base pairing in free duplex DNA. *Nucleic Acids Res.* **2015**, *43*, 7769-7778.

[2] Wright, P. E.; Dyson, H. J. Intrinsically unstructured proteins: Re-assessing the protein structure-function paradigm. *J. Mol. Biol.* **1999**, *293*, 321–331.

[3] Hoogsteen, K. Crystal and molecular structure of a hydrogen-bonded complex between 1-methylthymine and 9-methyladenine. *Acta Crystallogr.* **1963**, *16*, 907–916.

[4] Nikolova, E. N.; Zhou, H.; Gottardo, F. L.; Alvey, H. S.; Kimsey, I. J.; Al-Hashimi, H. M. A historical perspective on the role of Hoogsteen base pairs in nucleic acids. Biopolymers **2011**, *99*, 955–968.

[5] Shi, X.; Cui, W.; Pei, H. Hoogsteen base pairing in DNA-protein interactions and its role in genetic regulation. *Nat. Commun.* **2018**, *9*, 2361.

[6] Zhou, H.; Kimsey, I. J.; Nikolova, E. N.; Al-Hashimi, H. M. Hoogsteen base pairs: Expanding the repertoire of DNA recognition. *Chem. Rev.* **2019**, *119*, 7766–7786.

[7] Xu, Y.; McSally, J.; Andricioaei, I.; Al-Hashimi, H. M. Modulation of Hoogsteen dynamics on DNA recognition. *Nat. Commun.* **2018**, *9*, 1473.

[8] Nikan, M.; Shen, W.; Ma, H. Overcoming challenges in drugging structured nucleic acids: Hoogsteen base pairs as novel drug targets. *Trends Pharmacol. Sci.* **2016** *37*, 538–551.

[9] Aishima, J.; Gitti, R. K.; Noah, J. E.; Gan, H. H.; Schlick, T.; Wolberger, C. A Hoogsteen base pair embedded in undistorted B-DNA. *Nucleic Acids Res.* **2002**, *30*, 5244-5252.

[10] Oldfield, C. J.; Dunker, A. K. Intrinsically disordered proteins and intrinsically disordered protein regions. *Annu. Rev. Biochem.* **2014**, *83*, 553-584.

[11] Wright, P. E; Dyson, H. J. Intrinsically disordered proteins in cellular signalling and regulation. *Nat. Rev. Mol. Cell Biol.* **2015**, *16*, 18-29.

[12] Babu, M. M.; van der Lee, R.; de Groot, N. S.; Gsponer, J. Intrinsically disordered proteins: regulation and disease. *Curr. Opin. Struct. Biol.* **2011**, *21*, 432-440.

[13] Chiti, F.; Dobson, C. M. Protein misfolding, amyloid formation, and human disease: a summary of progress over the last decade. *Annu. Rev. Biochem.* **2017**, *86*, 27-68.

[14] Kole, K.; Gupta, A. M.; Chakrabarti, J. Conformational stability and order of Hoogsteen base pair induced by protein binding. *Biophys. Chem.* **2023**, *301*, 107079.

[15] Kole, K.; Chakrabarti, J. Binding of Homeodomain Proteins to DNA with Hoogsteen Base Pair. *J. Phys. Chem. B* **2025**, *129*, 1544-1554.

[16] Das, A.; Chakrabarti, J.; Ghosh, M. Conformational contribution to thermodynamics of binding in protein-peptide complexes through microscopic simulation. *Biophys. J.* **2013**, *104*, 1274-1284.

[17] Das, A.; Chakrabarti, J.; Ghosh, M. Conformational thermodynamics of metal-ion binding to a protein. *Chem. Phys. Lett.* **2013**, **581**, 91-95.

[18] Das, A.; Chakrabarti, J.; Ghosh, M. Thermodynamics of interfacial changes in a protein–protein complex. *Molecular Biosyst.* **2014**, *10*, 437-445.

[19] Sikdar, S.; Chakrabarti, J.; Ghosh, M. A microscopic insight from conformational thermodynamics to functional ligand binding in proteins. *Molecular Biosyst.* **2014**, *10*, 3280-3289.

[20] Mandal, S. C.; Maganti, L.; Mondal, M.; Chakrabarti, J. Microscopic insight to specificity of metal ion cofactor in DNA cleavage by restriction endonuclease EcoRV. *Biopolymers* **2020**, *111*, e23396.

[21] Moulick, A. G.; Chakrabarti, J. Fluctuation-Dominated Ligand Binding in Molten Globule Protein. J. Chem. Inf. Model. **2023**, **63**, 5583-5591.

[22] K¨astner, J. Umbrella sampling. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2011**, *1*, 932-942.

[23] Hub, J. S.; De Groot, B. L.; Van Der Spoel, D. g-wham-A free weighted histogram analysis implementation including robust error and autocorrelation estimates. *J. Chem. Theory Comput.* **2010**, *6*, 3713-3720.

[24] Lemkul, J. A.; Bevan, D. R. Assessing the stability of Alzheimer's amyloid protofibrils using molecular dynamics. *J. Phys. Chem. B* **2010**, *114*, 1652-1660.

[25] Barducci, A.; Bussi, G.; Parrinello, M. Well-tempered metadynamics: a smoothly converging and tunable free-energy method. *PRL* **2008**, *100*, 020603.

[26] Uversky, V. N. Alpha-synuclein misfolding and neurodegenerative diseases. *Curr. Protein Pept. Sci.* **2008**, *9*, 507–540.

[27] Asthana, S.; Bhattacharyya, D.; Kumari, S.; Nayak, P. S.; Saleem, M.; Bhunia, A.; Jha, S. Interaction with zinc oxide nanoparticle kinetically traps α-synuclein fibrillation into off-pathway non-toxic intermediates. *Int. J. Biol. Macromol.* **2020**, *150*, 68-79.

[28] Hashemi, M.; Lyubchenko, Y. L. Hybrid resolution molecular dynamics simulations of amyloid proteins interacting with membranes. *Methods* **2022**, *197*, 89-96.

[29] Cino, E. A.; Choy, W. Y.; Karttunen, M. Comparison of secondary structure formation using 10 different force fields in microsecond molecular dynamics simulations. J. Chem. Theory Comput. **2012**, *8*, 2725-2740.

[30] Monticelli, L.; Kandasamy, S. K.; Periole, X.; Larson, R. G.; Tieleman, D. P.; Marrink, S. J. The MARTINI coarse-grained force field: extension to proteins. *J. Chem. Theory Comput.* **2008**, *4*, 819-834.

[31] Trylska, J. Coarse-grained models to study dynamics of nanoscale biomolecules and their applications to the ribosome. J. Condens. Matter Phys. **2010**, *22*, 453101.

[32] Saunders, M. G.; Voth, G. A. Coarse-graining of multiprotein assemblies. *Curr. Opin. Struct. Biol.* **2012**, *22*, 144-150.

[33] Mukherjee, A.; Bagchi, B. Correlation between rate of folding, energy landscape, and topology in the folding of a model protein HP-36. *J. Chem. Phys.* **2003** *118*, 4733-4747.

[34] Kmiecik, S.; Gront, D.; Kolinski, M.; Wieteska, L.; Dawid, A. E.; Kolinski, A. Coarse-grained protein models and their applications. *Chem. Rev.* **2016**, *116*, 7898-7936.

[35] Detcheverry, F. A.; Pike, D. Q.; Nealey, P. F.; Müller, M.; de Pablo, J. J. Monte Carlo simulation of coarse grain polymeric systems. *PRL* **2009**, *102*, 197801.

[36] Mooney, C. Z. Monte carlo simulation. Sage **1997**, *116*.

[37] Kitayner, M.; Rozenberg, H.; Rohs, R.; Suad, O.; Rabinovich, D.; Honig, B.; Shakked, Z. Diversity in DNA recognition by p53 revealed by crystal structures with Hoogsteen base pairs. *Nat. Struct. Mol. Biol.* **2010**, *17*, 423-429.

[38] Patikoglou, G. A.; Kim, J. L.; Sun, L.; Yang, S. H.; Kodadek, T.; Burley, S. K. TATA element recognition by the TATA box-binding protein has been conserved throughout evolution. *Genes Dev.* **1999**, *13*, 3217-3230.

[39] Luscombe, N. M.; Laskowski, R. A.; Thornton, J. M. Amino acid–base interactions: a three-dimensional analysis of protein–DNA interactions at an atomic level. *Nucleic Acids Res.* **2001**, *29*, 2860-2874.

[40] Kalodimos, C. G.; Biris, N.; Bonvin, A. M.; Levandoski, M. M.; Guennuegues, M.; Boelens, R.; Kaptein, R. Structure and flexibility adaptation in nonspecific and specific protein-DNA complexes. *Science* **2004**, *305*, 386-389.

[41] Kalodimos, C. G.; Boelens, R.; Kaptein, R. A residue-specific view of the association and dissociation pathway in protein–DNA recognition. *Nat. Struct. Biol.* **2002**, *9*, 193-197.

[42] Yonetani, Y.; Kono, H. Dissociation free-energy profiles of specific and nonspecific DNA–protein complexes. *J. Phys. Chem. B* **2013**, *117*, 7535-7545.

[43] Jakubec, D.; Vondrasek, J. Can All-Atom Molecular Dynamics Simulations Quantitatively Describe Homeodomain–DNA Binding Equilibria? *J. Chem. Theory Comput.* **2019**, *15*, 2635-2648.

[44] Kumar, N.; Garg, P. Probing the molecular basis of cofactor affinity and conformational dynamics of mycobacterium tuberculosis elongation factor Tu: an integrated approach employing steered molecular dynamics and umbrella sampling simulations. *J. Phys. Chem. B* **2022**, *126*, 1447-1461.

[45] Ngo, S. T.; Vu, K. B.; Bui, L. M.; Vu, V. V. Effective estimation of ligand-binding affinity using biased sampling method. *ACS omega* **2019**, *4*, 3887-3893.

[46] Umbrella Sampling - GROMACS 2024.0 Tutorial - MD Tutorials (http://www.mdtutorials.com/gmx/umbrella/index.html)

[47] Chen, J.; Wu, M.; Gong, L.; Zhang, J.; Yan, B.; Liu, J.; Zhang, H.; Thundat, T.; Zeng, H. Mechanistic understanding and nanomechanics of multiple hydrogen-bonding interactions in aqueous environment. *J. Phys. Chem. C* **2019**, *123*, 4540-4548.

[48] Liu, B.; Li, J.; Qi, C.; Li, X.; Mai, T.; Zhang, J. Mechanism of asphaltene aggregation induced by supercritical CO 2: insights from molecular dynamics simulation. *RSC Adv.* **2017**, *7*, 50786-50793.

[49] Kumar, P.; Bhardwaj, V. K.; Purohit, R. Unveiling the intricate supramolecular chemistry of Y cyclodextrin epigallocatechin gallate inclusion complexes. *Ind. Eng. Chem. Res.* **2024**, *63*, 2544-2554.

[50] Lemkul, J. A.; Bevan, D. R. Assessing the stability of Alzheimer's amyloid protofibrils using molecular dynamics. *J. Phys. Chem. B* **2010**, *114*, 1652-1660.

[51] Cohen, I.; Huang, Y.; Chen, J.; Benesty, J.; Benesty, J.; Chen, J.; Huang, Y.; Cohen, I. Pearson correlation coefficient. *Noise reduction in speech processing* (**2009**) 1-4.

[52] Moulick, A. G.; Chakrabarti, J. Correlated dipolar and dihedral fluctuations in a protein. *Chem. Phy. Lett.* **2022**, *797*, 139574.

[53] Ratner, B. The correlation coefficient: Its values range between +1/-1, or do they? *Journal of targeting, measurement and analysis for marketing* **2009**, *17*, 139-142.

[54] Yang, C.; Kim, E.; Lim, M.; Pak, Y. Computational probing of Watson–crick/hoogsteen breathing in a DNA duplex containing N1-methylated adenine. *J. Chem. Theory Comput.* **2018**, *15*, 751-761.

[55] Stone, S. E.; Ray, D.; Andricioaei, I. Force-field-dependent DNA breathing dynamics: a case study of Hoogsteen Base pairing in A6-DNA. *J. Chem. Inf. Model.* **2022**, *62*, 6749-6761.

[56] Torrie, G. M.; Valleau, J. P. Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. . *Comput. Phys.* **1977**, *23*, 187-199.

[57] Laio, A.; Parrinello, M. Escaping free-energy minima. *Proceedings of the national academy of sciences* **2002** *99*, 12562-12566.

[58] Barducci, A.; Bussi, G.; Parrinello, M. Well-tempered metadynamics: a smoothly converging and tunable free-energy method. *PRL* **2008**, *100*, 020603.

[59] Fu, H.; Zhang, H.; Chen, H.; Shao, X.; Chipot, C.; Cai, W. Zooming across the free-energy landscape: shaving barriers, and flooding valleys. *J. Phys. Chem. Lett.* **2018**, *9*, 4738-4745.

[60] Barducci, A.; Bussi, G.; Parrinello, M. Well-tempered metadynamics: a smoothly converging and tunable free-energy method. *PRL* **2008**, *100*, 020603.

[61] Nikolova, E. N.; Kim, E.; Wise,A. A.; O'Brien, P. J.; Andricioaei, I.; Al-Hashimi, H. M. Transient Hoogsteen base pairs in canonical duplex DNA. *Nature* **2011**, *470*, 498–502.

[62] Chiti, F.; Dobson, C. M. Protein misfolding, functional amyloid, and human disease. *Annu. Rev. Biochem.* **2006**, *75*, 333-366.

[63] Chiti, F.; Dobson, C. M. Protein misfolding, amyloid formation, and human disease: a summary of progress over the last decade. *Annu. Rev. Biochem.* **2017**, *86*, 27-68.

[64] Eisenberg, D. S.; Sawaya, M. R. Structural studies of amyloid proteins at the molecular level. *Annu. Rev. Biochem.* **2017**, 86, 69-95.

[65] Giasson, B. I.; Jakes, R.; Goedert, M.; Duda, J. E.; Leight, S.; Trojanowski, J. Q.; Lee, V. M. A panel of epitope-specific antibodies detects protein domains distributed throughout human α-synuclein in lewy bodies of Parkinson's disease. *J. Neurosci. Res.* **2000**, *59*, 528-533.

[66] Spillantini, M. G.; Schmidt, M. L.; Lee, V. M. Y.; Trojanowski, J. Q.; Jakes, R.; Goedert, M. α-Synuclein in Lewy bodies. *Nature* **1997**, *388*, 839-840.

[67] Baba, M.; Nakajo, S.; Tu, P. H.; Tomita, T.; Nakaya, K.; Lee, V. M.; Trojanowski J. Q.; Iwatsubo, T. Aggregation of alpha-synuclein in Lewy bodies of sporadic Parkinson's disease and dementia with Lewy bodies. *AJP* **1998**, *152*, 879.

[68] Burre, J. The synaptic function of α-synuclein. *JPD* **2015**, *5*, 699-713.

[69] Yamauchi, M.; Okumura, H. Dimerization of α-Synuclein fragments studied by isothermal–isobaric replica-permutation molecular dynamics simulation. *J. Chem. Inf. Model.* **2021**, *61*, 1307-1321.

[70] Kamelabad, M. R.; Sardroodi, J. J.; Ebrahimzadeh, A. R.; Ajamgard, M. Influence of curcumin and rosmarinic acid on disrupting the general properties of Alpha-Synuclein oligomer: Molecular dynamics simulation. *J. Mol. Graph. Model.* **2021**, *107*, 107963.

[71] Sittel, F.; Stock, G. Perspective: Identification of collective variables and metastable states of protein dynamics. *J. Chem. Phys.* **2018**, *149*.

[72] Trylska, J. Coarse-grained models to study dynamics of nanoscale biomolecules and their applications to the ribosome. *J. Condens. Matter Phys.* **2010**, *22*, 453101.

[73] Saunders, M. G.; Voth, G. A. Coarse-graining of multiprotein assemblies. *Curr. Opin. Struct. Biol.* **2012**, *22*, 144-150.

[74] Saunders, M. G.; Voth, G. A. Coarse-graining methods for computational biology. *Annu. Rev. Biophys.* **2013**, *42*, 73-93.

[75] Ulmer, T. S.; Ramirez, B. E.; Delaglio, F.; Bax, A. Evaluation of backbone proton positions and dynamics in a small protein by liquid crystal NMR spectroscopy. *J. Am. Chem. Soc.* **2003**, *125*, 9179-9191.